

TMPWG Guidance Document On Interpreting and Using Emissions Databases Containing Non-detection Values

The ICCR Testing and Monitoring Protocols Workgroup (TMPWG) reviewed ICCR issues which result from emission testing that produce reports of non-detection concentrations and we formulated procedures and recommendations to deal with such reports. With existing EPA ICCR and other databases, we assumed that there is a need to obtain mean and/or median values and variability for data sets for various reasons. The most critical reason might be determining whether a toxic emission from a group of potential emissions sources warrants further consideration in the ICCR process.

We believe that any decision to control HAP emissions from combustion sources should be made on the basis of fuel composition, combustion science, and actual observations. No decisions leading to the imposition of control devices or emission limits on combustion processes should be made that are based on emission levels derived from default HAP concentrations calculated from method detection levels.

We strongly encourage all Source Work Groups (SWG) to define a "critical concentration level" below which HAP emissions are not significant for the purposes of data gathering as part their planning future emissions testing related to the ICCR process. The TMPWG would be available to assist in designing a test protocol to reach that goal.

Our recommendation is for the SWG to follow these steps when making decisions that involve interpretation of databases with reported non-detection reports for some pollutants:

Step 1: In addition to emission concentration levels, consider fuel composition, scientific and engineering data to focus efforts towards HAP emissions that are potentially significant. Doing this is critical for non-detection and other issues because it is impractical and simply not necessary to solely rely on stack testing to rule out all 189 chemicals that appear on the EPA toxics substances list. Material balances using fuel composition and consumption rates along with published flame chemistry science are most useful in this regard. We believe that most potential non-detection issues can best be circumvented by completion of Step 1.

Step 2: Associate detection limits with individual source tests that resulted in non-detection reports either by retrieving the detection limit from the database, or assigning conservative detection limits based on the descriptions of the measurement procedures. Retrieval is preferred, but assignments without uncertainties are not realistic.

Step 3: Assume that the actual concentrations could be as high as the detection limits to determine if the emissions have the potential to be important in the ICCR process even with the highest potential concentration. If the answer is no, there is no need for the following steps.

Step 4: Create a data subset which contains only those source measurements that have the lowest detection limits, and repeat Step 3. If the answer becomes no, and if the reviewers agree that the subset is representative of the industry emissions sources, there is no need for the following steps.

Step 5: Define the detection level that is needed to resolve the ICCR issues relative to the emission of a specific HAP, this sets a "critical concentration level" below which emissions are not significant for the most restrictive ICCR issues.

Step 6: The best course of action for filling data gaps is the collection of data using methodology with the appropriately low detection limits. A less desirable alternative is to use the 1/2 detection limit substitution method on an existing database.

The publications by Helsel¹, Coleman, *et. al.*², and Zorn *et. al.*³ address the issues of dealing with databases that contain a mixture of detection and non-detection values, and give procedures for determining mean and median values. Procedures that are discussed and provided range from simple substitution to complex statistical methods. The publications show mean and median values can be generated with the highest certainty when:

1. The detection limits for each measurement is known with certainty, and when the detection limit and the definition of the detection limit are consistent.
2. The ratio of non-detection to detection values is less than 1, and when there are enough detectable values so the mean and median values are not dominated by statistical outliers.

Simple substitution methods using the ½ detection limits generally perform poorly as compared to the more complex statistical methodology when the above conditions were met. Substitution of zero for the detection limit was discouraged because it will result in a low bias, and substitution of the detection limit was discouraged because it will result in a high bias. We believe that limitations in the existing database make the ½ detection limits substitution method the most applicable for working with existing databases. But testing with appropriate detection limits is the most reliable approach and is our first recommendation for filling data gaps.

¹ Dennis R. Helsel, "Less than obvious: statistical treatment of data below the detection limit", Environ. Sci. Technol., 1990, Vol. 24, pp. 1766 - 1774.

² David Coleman, *et.al.*, "Regulation between detection limits, quantification limits, and significant digits" Chemometrics and Intelligent Laboratory Systems, 1997, Vol. 37, pp 71-80.

³ Michael E. Zorn, *et. al.*, "Weighted least-squares approach to calculating limits of detection and quantification by modeling variability as a function of concentration", Anal. Chem., 1997, Vol. 69, pp. 3069-3075.